wwPDB X-ray Structure Validation Summary Report (i)

Dec 14, 2016 – 10:54 AM EST

PDB ID	:	5CDQ
Title	:	2.95A structure of Moxifloxacin with S.aureus DNA gyrase and DNA
Authors	:	Bax, B.D.; Srikannathasan, V.; Chan, P.F.
Deposited on	:	2015-07-04
Resolution	:	2.95 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

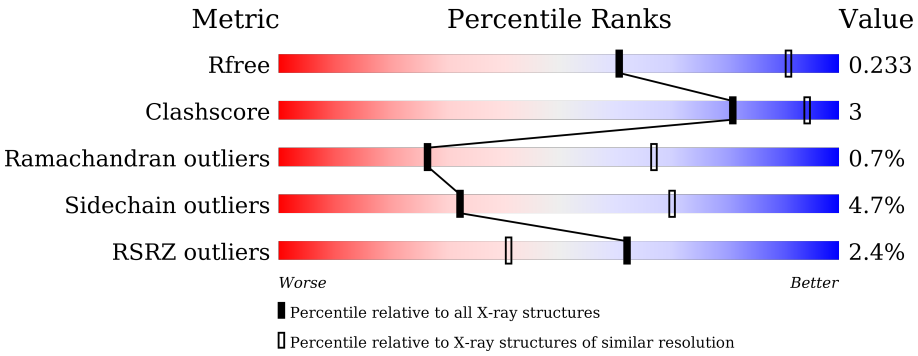
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.95 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2184 (3.00-2.92)
Clashscore	102246	2552 (3.00-2.92)
Ramachandran outliers	100387	2468 (3.00-2.92)
Sidechain outliers	100360	2471 (3.00-2.92)
RSRZ outliers	91569	2201 (3.00-2.92)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	481	90% 9%
1	C	481	91% 8%
1	R	481	3% 89% 11%
1	T	481	90% 10%
2	B	193	4% 87% 12%
2	D	193	7% 78% 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	S	193	4% 83% 16%
2	U	193	6% 81% 15%
3	E	20	90% 10%
3	F	20	85% 15%
3	V	20	5% 75% 20% 5%
3	W	20	80% 15% 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	GOL	A	504	-	-	-	X
5	GOL	C	504	-	-	-	X
5	GOL	R	4303	-	-	-	X
5	GOL	R	4304	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 23073 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA gyrase subunit A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	481	Total	C	N	O	P	S	0	2	0
			3767	2343	680	727	1	16			
1	C	481	Total	C	N	O	P	S	0	2	0
			3790	2357	689	727	1	16			
1	R	481	Total	C	N	O	P	S	0	2	0
			3804	2366	692	729	1	16			
1	T	481	Total	C	N	O	P	S	0	2	0
			3764	2343	679	726	1	15			

- Molecule 2 is a protein called DNA gyrase subunit B,DNA gyrase subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	191	Total	C	N	O	S	0	0	0
			1468	923	251	285	9			
2	D	190	Total	C	N	O	S	0	0	0
			1429	895	248	279	7			
2	S	193	Total	C	N	O	S	0	0	0
			1499	943	260	287	9			
2	U	192	Total	C	N	O	S	0	0	0
			1454	915	251	279	9			

There are 144 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	LEU	deletion	UNP P66937
B	?	-	TYR	deletion	UNP P66937
B	?	-	LYS	deletion	UNP P66937
B	?	-	LEU	deletion	UNP P66937
B	?	-	THR	deletion	UNP P66937
B	?	-	GLN	deletion	UNP P66937
B	?	-	GLY	deletion	UNP P66937
B	?	-	LYS	deletion	UNP P66937

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	GLN	deletion	UNP P66937
B	?	-	LYS	deletion	UNP P66937
B	?	-	TYR	deletion	UNP P66937
B	?	-	TYR	deletion	UNP P66937
B	?	-	VAL	deletion	UNP P66937
B	?	-	TYR	deletion	UNP P66937
B	?	-	ASN	deletion	UNP P66937
B	?	-	ASP	deletion	UNP P66937
B	?	-	ARG	deletion	UNP P66937
B	?	-	GLU	deletion	UNP P66937
B	?	-	LEU	deletion	UNP P66937
B	?	-	ASP	deletion	UNP P66937
B	?	-	LYS	deletion	UNP P66937
B	?	-	LEU	deletion	UNP P66937
B	?	-	LYS	deletion	UNP P66937
B	?	-	SER	deletion	UNP P66937
B	?	-	GLU	deletion	UNP P66937
B	?	-	LEU	deletion	UNP P66937
B	?	-	ASN	deletion	UNP P66937
B	?	-	PRO	deletion	UNP P66937
B	?	-	THR	deletion	UNP P66937
B	?	-	PRO	deletion	UNP P66937
B	?	-	LYS	deletion	UNP P66937
B	?	-	TRP	deletion	UNP P66937
B	?	-	SER	deletion	UNP P66937
B	?	-	ILE	deletion	UNP P66937
B	544	THR	ALA	linker	UNP P66937
B	545	GLY	ARG	linker	UNP P66937
D	?	-	LEU	deletion	UNP P66937
D	?	-	TYR	deletion	UNP P66937
D	?	-	LYS	deletion	UNP P66937
D	?	-	LEU	deletion	UNP P66937
D	?	-	THR	deletion	UNP P66937
D	?	-	GLN	deletion	UNP P66937
D	?	-	GLY	deletion	UNP P66937
D	?	-	LYS	deletion	UNP P66937
D	?	-	GLN	deletion	UNP P66937
D	?	-	LYS	deletion	UNP P66937
D	?	-	TYR	deletion	UNP P66937
D	?	-	TYR	deletion	UNP P66937
D	?	-	VAL	deletion	UNP P66937
D	?	-	TYR	deletion	UNP P66937

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	?	-	ASN	deletion	UNP P66937
D	?	-	ASP	deletion	UNP P66937
D	?	-	ARG	deletion	UNP P66937
D	?	-	GLU	deletion	UNP P66937
D	?	-	LEU	deletion	UNP P66937
D	?	-	ASP	deletion	UNP P66937
D	?	-	LYS	deletion	UNP P66937
D	?	-	LEU	deletion	UNP P66937
D	?	-	LYS	deletion	UNP P66937
D	?	-	SER	deletion	UNP P66937
D	?	-	GLU	deletion	UNP P66937
D	?	-	LEU	deletion	UNP P66937
D	?	-	ASN	deletion	UNP P66937
D	?	-	PRO	deletion	UNP P66937
D	?	-	THR	deletion	UNP P66937
D	?	-	PRO	deletion	UNP P66937
D	?	-	LYS	deletion	UNP P66937
D	?	-	TRP	deletion	UNP P66937
D	?	-	SER	deletion	UNP P66937
D	?	-	ILE	deletion	UNP P66937
D	544	THR	ALA	linker	UNP P66937
D	545	GLY	ARG	linker	UNP P66937
S	?	-	LEU	deletion	UNP P66937
S	?	-	TYR	deletion	UNP P66937
S	?	-	LYS	deletion	UNP P66937
S	?	-	LEU	deletion	UNP P66937
S	?	-	THR	deletion	UNP P66937
S	?	-	GLN	deletion	UNP P66937
S	?	-	GLY	deletion	UNP P66937
S	?	-	LYS	deletion	UNP P66937
S	?	-	GLN	deletion	UNP P66937
S	?	-	LYS	deletion	UNP P66937
S	?	-	TYR	deletion	UNP P66937
S	?	-	TYR	deletion	UNP P66937
S	?	-	VAL	deletion	UNP P66937
S	?	-	TYR	deletion	UNP P66937
S	?	-	ASN	deletion	UNP P66937
S	?	-	ASP	deletion	UNP P66937
S	?	-	ARG	deletion	UNP P66937
S	?	-	GLU	deletion	UNP P66937
S	?	-	LEU	deletion	UNP P66937
S	?	-	ASP	deletion	UNP P66937

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
S	?	-	LYS	deletion	UNP P66937
S	?	-	LEU	deletion	UNP P66937
S	?	-	LYS	deletion	UNP P66937
S	?	-	SER	deletion	UNP P66937
S	?	-	GLU	deletion	UNP P66937
S	?	-	LEU	deletion	UNP P66937
S	?	-	ASN	deletion	UNP P66937
S	?	-	PRO	deletion	UNP P66937
S	?	-	THR	deletion	UNP P66937
S	?	-	PRO	deletion	UNP P66937
S	?	-	LYS	deletion	UNP P66937
S	?	-	TRP	deletion	UNP P66937
S	?	-	SER	deletion	UNP P66937
S	?	-	ILE	deletion	UNP P66937
S	544	THR	ALA	linker	UNP P66937
S	545	GLY	ARG	linker	UNP P66937
U	?	-	LEU	deletion	UNP P66937
U	?	-	TYR	deletion	UNP P66937
U	?	-	LYS	deletion	UNP P66937
U	?	-	LEU	deletion	UNP P66937
U	?	-	THR	deletion	UNP P66937
U	?	-	GLN	deletion	UNP P66937
U	?	-	GLY	deletion	UNP P66937
U	?	-	LYS	deletion	UNP P66937
U	?	-	GLN	deletion	UNP P66937
U	?	-	LYS	deletion	UNP P66937
U	?	-	TYR	deletion	UNP P66937
U	?	-	TYR	deletion	UNP P66937
U	?	-	VAL	deletion	UNP P66937
U	?	-	TYR	deletion	UNP P66937
U	?	-	ASN	deletion	UNP P66937
U	?	-	ASP	deletion	UNP P66937
U	?	-	ARG	deletion	UNP P66937
U	?	-	GLU	deletion	UNP P66937
U	?	-	LEU	deletion	UNP P66937
U	?	-	ASP	deletion	UNP P66937
U	?	-	LYS	deletion	UNP P66937
U	?	-	LEU	deletion	UNP P66937
U	?	-	LYS	deletion	UNP P66937
U	?	-	SER	deletion	UNP P66937
U	?	-	GLU	deletion	UNP P66937
U	?	-	LEU	deletion	UNP P66937

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
U	?	-	ASN	deletion	UNP P66937
U	?	-	PRO	deletion	UNP P66937
U	?	-	THR	deletion	UNP P66937
U	?	-	PRO	deletion	UNP P66937
U	?	-	LYS	deletion	UNP P66937
U	?	-	TRP	deletion	UNP P66937
U	?	-	SER	deletion	UNP P66937
U	?	-	ILE	deletion	UNP P66937
U	544	THR	ALA	linker	UNP P66937
U	545	GLY	ARG	linker	UNP P66937

- Molecule 3 is a DNA chain called DNA (5'-D(P*GP*AP*GP*CP*GP*TP*AP*T*GP*GP*CP*CP*AP*TP*AP*CP*GP*CP*TP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	20	Total	C	N	O	P	0	0	0
			392	186	73	114	19			
3	F	20	Total	C	N	O	P	0	0	0
			389	187	73	111	18			
3	V	20	Total	C	N	O	P	0	0	0
			392	186	73	114	19			
3	W	20	Total	C	N	O	P	0	0	0
			391	188	73	112	18			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

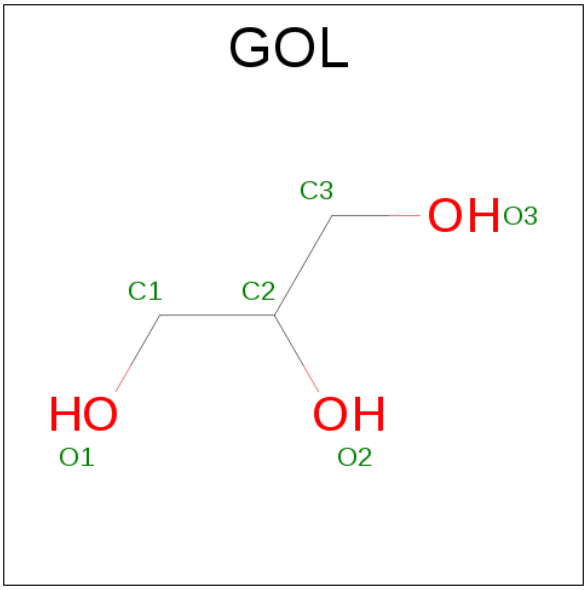
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	1	Total	Mg	0	0
			1	1		
4	B	1	Total	Mg	0	0
			1	1		
4	C	2	Total	Mg	0	0
			2	2		
4	W	2	Total	Mg	0	0
			2	2		
4	A	2	Total	Mg	0	0
			2	2		
4	T	2	Total	Mg	0	0
			2	2		
4	U	1	Total	Mg	0	0
			1	1		

Continued on next page...

Continued from previous page...

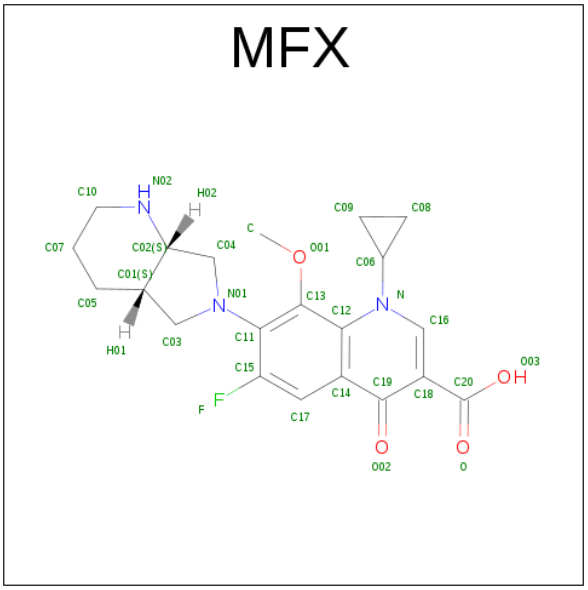
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	R	1	Total Mg 1 1	0	0
4	S	1	Total Mg 1 1	0	0
4	F	1	Total Mg 1 1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	R	1	Total C O 6 3 3	0	0
5	R	1	Total C O 6 3 3	0	0
5	R	1	Total C O 6 3 3	0	0
5	T	1	Total C O 6 3 3	0	0

- Molecule 6 is 1-cyclopropyl-6-fluoro-8-methoxy-7-[(4aS,7aS)-octahydro-6H-pyrrolo[3,4-b]pyridin-6-yl]-4-oxo-1,4-dihydroquinoline-3-carboxylic acid (three-letter code: MFX) (formula: $C_{21}H_{24}FN_3O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	E	1	Total C F N O 29 21 1 3 4	0	0
6	F	1	Total C F N O 29 21 1 3 4	0	0
6	F	1	Total C F N O 29 21 1 3 4	0	0
6	V	1	Total C F N O 29 21 1 3 4	0	0
6	W	1	Total C F N O 29 21 1 3 4	0	0
6	W	1	Total C F N O 29 21 1 3 4	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	42	Total O 42 42	0	0
7	B	21	Total O 21 21	0	0
7	C	41	Total O 41 41	0	0
7	D	5	Total O 5 5	0	0

Continued on next page...

Continued from previous page...

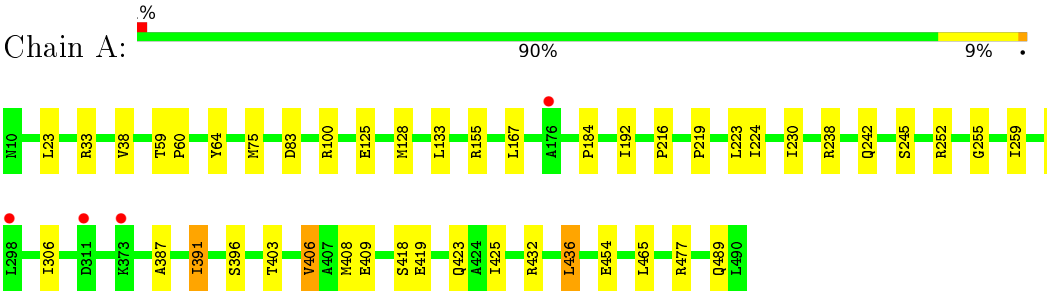
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	E	21	Total O 21 21	0	0
7	F	17	Total O 17 17	0	0
7	R	59	Total O 59 59	0	0
7	S	20	Total O 20 20	0	0
7	T	37	Total O 37 37	0	0
7	U	10	Total O 10 10	0	0
7	V	12	Total O 12 12	0	0
7	W	13	Total O 13 13	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

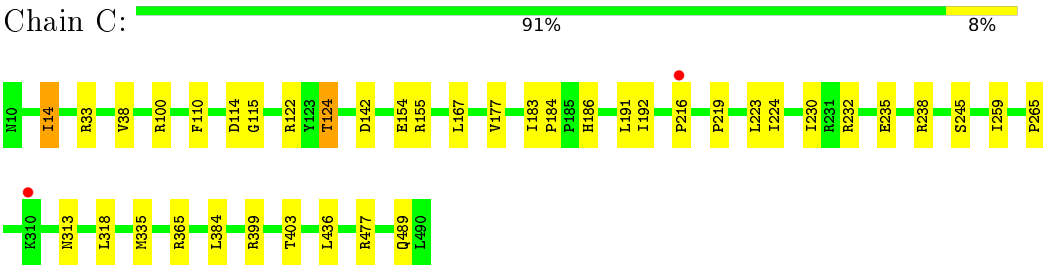
- Molecule 1: DNA gyrase subunit A

Chain A:



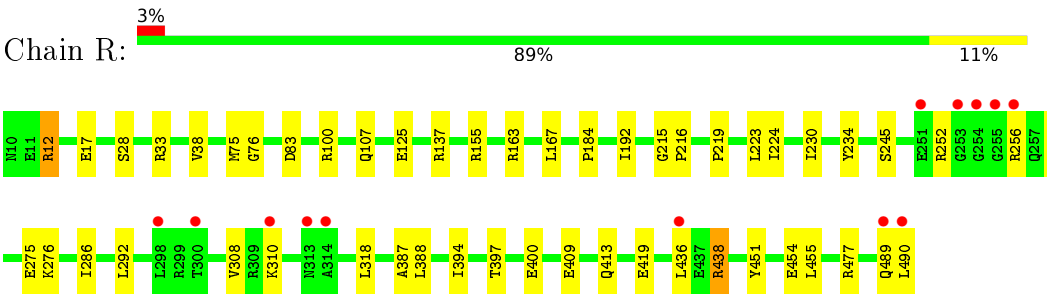
- Molecule 1: DNA gyrase subunit A

Chain C:



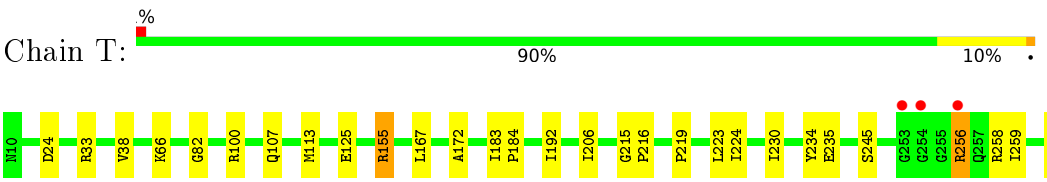
- Molecule 1: DNA gyrase subunit A

Chain R:

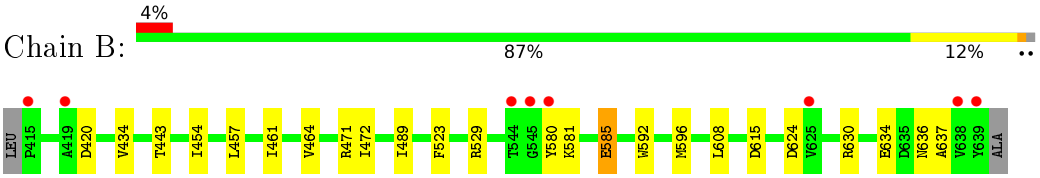


- Molecule 1: DNA gyrase subunit A

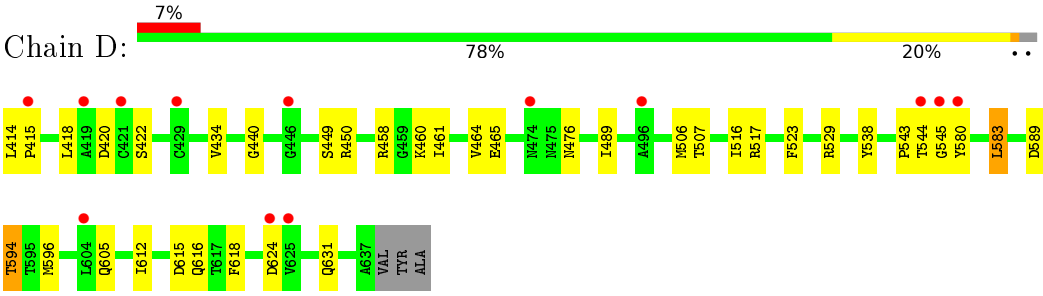
Chain T:



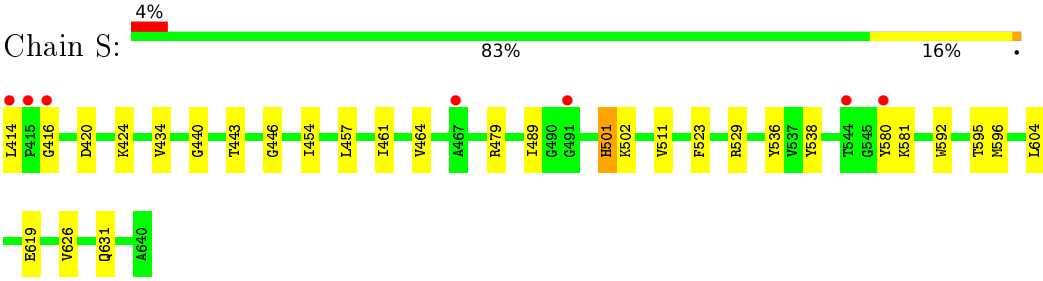
• Molecule 2: DNA gyrase subunit B,DNA gyrase subunit B



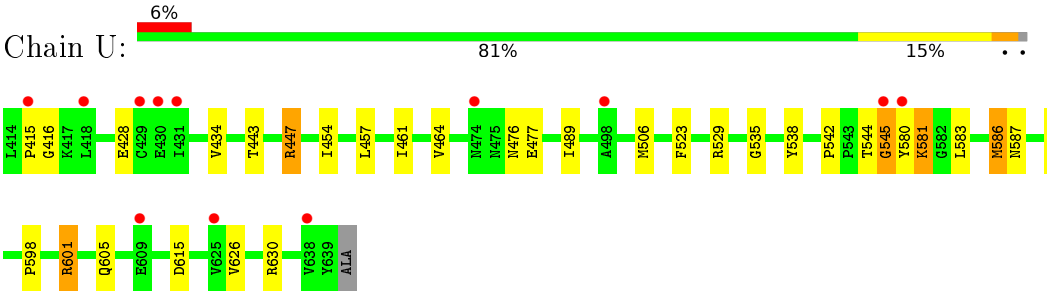
• Molecule 2: DNA gyrase subunit B,DNA gyrase subunit B



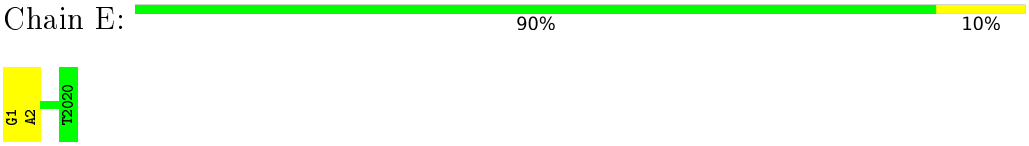
• Molecule 2: DNA gyrase subunit B,DNA gyrase subunit B



• Molecule 2: DNA gyrase subunit B,DNA gyrase subunit B

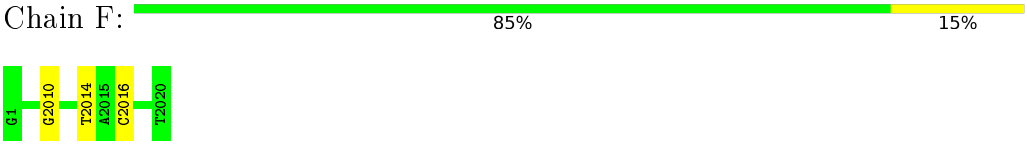


• Molecule 3: DNA (5'-D(P*GP*AP*GP*CP*GP*TP*AP*T*GP*GP*CP*CP*AP*TP*AP*CP*GP*CP*TP*T)-3')



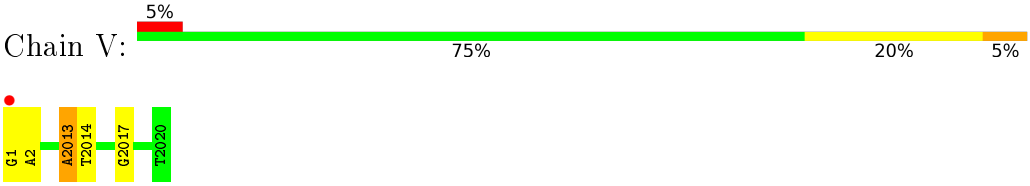
• Molecule 3: DNA (5'-D(P*GP*AP*GP*CP*GP*TP*AP*T*GP*GP*CP*CP*AP*TP*AP*CP*GP*CP*TP*T)-3')

Chain F:



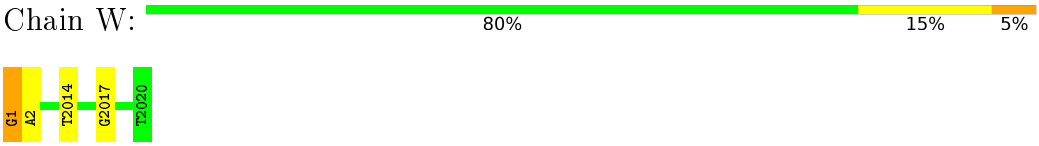
• Molecule 3: DNA (5'-D(P*GP*AP*GP*CP*GP*TP*AP*T*GP*GP*CP*CP*AP*TP*AP*CP*GP*CP*TP*T)-3')

Chain V:



• Molecule 3: DNA (5'-D(P*GP*AP*GP*CP*GP*TP*AP*T*GP*GP*CP*CP*AP*TP*AP*CP*GP*CP*TP*T)-3')

Chain W:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	87.93Å 170.55Å 125.67Å 90.00° 103.30° 90.00°	Depositor
Resolution (Å)	19.99 – 2.95 19.99 – 2.95	Depositor EDS
% Data completeness (in resolution range)	98.8 (19.99-2.95) 98.8 (19.99-2.95)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.25 (at 2.93Å)	Xtriage
Refinement program	BUSTER-TNT	Depositor
R, R_{free}	0.174 , 0.218 0.190 , 0.233	Depositor DCC
R_{free} test set	3795 reflections (5.36%)	DCC
Wilson B-factor (Å^2)	67.6	Xtriage
Anisotropy	0.858	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 61.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	23073	wwPDB-VP
Average B, all atoms (Å^2)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.14% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MFX, GOL, MG, PTR

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.41	0/3796	0.62	0/5123
1	C	0.42	0/3820	0.62	0/5154
1	R	0.41	0/3834	0.61	0/5168
1	T	0.41	0/3794	0.62	0/5122
2	B	0.44	0/1492	0.66	0/2017
2	D	0.43	0/1452	0.67	0/1969
2	S	0.44	0/1524	0.65	0/2059
2	U	0.46	0/1478	0.67	0/2003
3	E	1.04	0/438	0.94	0/673
3	F	1.00	1/435 (0.2%)	0.96	1/669 (0.1%)
3	V	1.00	0/438	0.97	2/673 (0.3%)
3	W	1.05	1/437 (0.2%)	1.00	1/672 (0.1%)
All	All	0.50	2/22938 (0.0%)	0.67	4/31302 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	W	1	DG	C3'-O3'	6.79	1.52	1.44
3	F	2010	DG	O5'-C5'	-5.56	1.28	1.42

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	V	2014	DT	O4'-C4'-C3'	-6.10	102.06	104.50
3	V	2013	DA	O4'-C1'-N9	5.34	111.74	108.00
3	W	2014	DT	O4'-C4'-C3'	-5.27	102.39	104.50
3	F	2014	DT	O4'-C4'-C3'	-5.08	102.47	104.50

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3767	0	3745	22	0
1	C	3790	0	3790	21	0
1	R	3804	0	3826	28	0
1	T	3764	0	3745	22	0
2	B	1468	0	1423	10	0
2	D	1429	0	1353	17	0
2	S	1499	0	1473	15	0
2	U	1454	0	1400	14	0
3	E	392	0	212	1	0
3	F	389	0	212	1	0
3	V	392	0	212	4	0
3	W	391	0	213	3	0
4	A	2	0	0	0	0
4	B	1	0	0	0	0
4	C	2	0	0	0	0
4	D	1	0	0	0	0
4	F	1	0	0	0	0
4	R	1	0	0	0	0
4	S	1	0	0	0	0
4	T	2	0	0	0	0
4	U	1	0	0	0	0
4	W	2	0	0	0	0
5	A	12	0	16	0	0
5	C	12	0	16	1	0
5	R	18	0	24	0	0
5	T	6	0	8	1	0
6	E	29	0	23	1	0
6	F	58	0	46	2	0
6	V	29	0	23	3	0
6	W	58	0	46	0	0
7	A	42	0	0	0	0
7	B	21	0	0	0	0
7	C	41	0	0	0	0
7	D	5	0	0	0	0
7	E	21	0	0	0	0
7	F	17	0	0	0	0
7	R	59	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	S	20	0	0	0	0
7	T	37	0	0	1	0
7	U	10	0	0	0	0
7	V	12	0	0	0	0
7	W	13	0	0	0	0
All	All	23073	0	21806	145	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

The worst 5 of 145 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:T:313:ASN:HD22	1:T:316:VAL:HG23	1.41	0.85
2:D:460:LYS:HA	2:D:516:ILE:HD11	1.59	0.82
2:D:592:TRP:HA	2:D:596:MET:HB2	1.69	0.73
2:U:587:ASN:HB3	2:U:590:GLN:HB2	1.72	0.72
1:R:76:GLY:HA2	1:T:66[B]:LYS:HG3	1.72	0.71

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	480/481 (100%)	462 (96%)	15 (3%)	3 (1%)	30 70
1	C	480/481 (100%)	464 (97%)	14 (3%)	2 (0%)	39 78
1	R	480/481 (100%)	465 (97%)	14 (3%)	1 (0%)	52 86
1	T	480/481 (100%)	463 (96%)	15 (3%)	2 (0%)	39 78
2	B	189/193 (98%)	183 (97%)	4 (2%)	2 (1%)	17 56

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	D	188/193 (97%)	178 (95%)	7 (4%)	3 (2%)	12 45
2	S	191/193 (99%)	182 (95%)	7 (4%)	2 (1%)	19 58
2	U	190/193 (98%)	180 (95%)	5 (3%)	5 (3%)	7 30
All	All	2678/2696 (99%)	2577 (96%)	81 (3%)	20 (1%)	26 67

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	637	ALA
2	D	545	GLY
2	D	580	TYR
2	U	544	THR
1	A	33	ARG

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	397/415 (96%)	383 (96%)	14 (4%)	43 78
1	C	401/415 (97%)	391 (98%)	10 (2%)	55 85
1	R	405/415 (98%)	392 (97%)	13 (3%)	46 80
1	T	396/415 (95%)	381 (96%)	15 (4%)	40 76
2	B	151/160 (94%)	143 (95%)	8 (5%)	28 65
2	D	141/160 (88%)	124 (88%)	17 (12%)	6 23
2	S	156/160 (98%)	143 (92%)	13 (8%)	14 43
2	U	147/160 (92%)	135 (92%)	12 (8%)	14 44
All	All	2194/2300 (95%)	2092 (95%)	102 (5%)	32 71

5 of 102 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	616	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	R	419	GLU
2	U	489	ILE
2	D	624	ASP
1	R	275	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	605	GLN
1	R	257	GLN
1	T	354	HIS
2	D	501	HIS
2	D	597	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

4 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	PTR	A	123	1,3	13,16,17	1.69	2 (15%)	19,22,24	1.60	5 (26%)
1	PTR	C	123	1,3	13,16,17	0.97	0	19,22,24	1.47	3 (15%)
1	PTR	R	123	1,3	13,16,17	1.21	1 (7%)	19,22,24	1.33	4 (21%)
1	PTR	T	123	1,3	13,16,17	0.85	0	19,22,24	1.40	3 (15%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means

no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	PTR	A	123	1,3	-	0/9/11/13	0/1/1/1
1	PTR	C	123	1,3	-	0/9/11/13	0/1/1/1
1	PTR	R	123	1,3	-	0/9/11/13	0/1/1/1
1	PTR	T	123	1,3	-	0/9/11/13	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	123	PTR	CE1-CD1	2.35	1.43	1.38
1	R	123	PTR	CE2-CZ	2.73	1.44	1.38
1	A	123	PTR	CE1-CZ	4.33	1.47	1.38

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	123	PTR	O-C-CA	-3.60	116.08	125.72
1	T	123	PTR	P-OH-CZ	-2.74	116.07	123.85
1	A	123	PTR	OH-CZ-CE2	-2.68	110.86	119.22
1	T	123	PTR	O-C-CA	-2.56	118.85	125.72
1	C	123	PTR	P-OH-CZ	-2.49	116.78	123.85

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 28 ligands modelled in this entry, 14 are monoatomic - leaving 14 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected

value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	GOL	A	503	-	5,5,5	0.11	0	5,5,5	0.27	0
5	GOL	A	504	-	5,5,5	0.13	0	5,5,5	0.17	0
5	GOL	C	503	-	5,5,5	0.12	0	5,5,5	0.19	0
5	GOL	C	504	-	5,5,5	0.15	0	5,5,5	0.43	0
6	MFX	E	2101	4	25,33,33	3.40	7 (28%)	32,50,50	3.90	15 (46%)
6	MFX	F	2101	4	25,33,33	3.14	7 (28%)	32,50,50	3.83	17 (53%)
6	MFX	F	2103	4	25,33,33	3.41	5 (20%)	32,50,50	3.86	12 (37%)
5	GOL	R	4302	-	5,5,5	0.16	0	5,5,5	0.36	0
5	GOL	R	4303	-	5,5,5	0.11	0	5,5,5	0.31	0
5	GOL	R	4304	-	5,5,5	0.12	0	5,5,5	0.21	0
5	GOL	T	4503	-	5,5,5	0.13	0	5,5,5	0.24	0
6	MFX	V	2101	4	25,33,33	3.38	6 (24%)	32,50,50	3.82	15 (46%)
6	MFX	W	2101	4	25,33,33	3.12	5 (20%)	32,50,50	3.83	13 (40%)
6	MFX	W	2103	4	25,33,33	3.44	6 (24%)	32,50,50	3.51	11 (34%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GOL	A	503	-	-	0/4/4/4	0/0/0/0
5	GOL	A	504	-	-	0/4/4/4	0/0/0/0
5	GOL	C	503	-	-	0/4/4/4	0/0/0/0
5	GOL	C	504	-	-	0/4/4/4	0/0/0/0
6	MFX	E	2101	4	-	0/6/35/35	0/4/5/5
6	MFX	F	2101	4	-	0/6/35/35	0/4/5/5
6	MFX	F	2103	4	-	0/6/35/35	0/4/5/5
5	GOL	R	4302	-	-	0/4/4/4	0/0/0/0
5	GOL	R	4303	-	-	0/4/4/4	0/0/0/0
5	GOL	R	4304	-	-	0/4/4/4	0/0/0/0
5	GOL	T	4503	-	-	0/4/4/4	0/0/0/0
6	MFX	V	2101	4	-	0/6/35/35	0/4/5/5
6	MFX	W	2101	4	-	0/6/35/35	0/4/5/5
6	MFX	W	2103	4	-	0/6/35/35	0/4/5/5

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	V	2101	MFX	C13-C12	2.03	1.51	1.43
6	E	2101	MFX	C08-C06	2.05	1.53	1.48
6	E	2101	MFX	C13-C12	2.09	1.51	1.43
6	F	2101	MFX	C13-C12	2.13	1.51	1.43
6	F	2101	MFX	C17-C15	2.16	1.39	1.35

The worst 5 of 83 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	F	2103	MFX	C04-N01-C11	-9.01	109.02	122.82
6	F	2103	MFX	C04-N01-C03	-7.35	102.27	112.21
6	W	2103	MFX	C03-N01-C11	-6.64	112.66	122.82
6	V	2101	MFX	C04-N01-C03	-6.31	103.69	112.21
6	E	2101	MFX	C03-N01-C11	-6.07	113.53	122.82

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	504	GOL	1	0
6	E	2101	MFX	1	0
6	F	2101	MFX	1	0
6	F	2103	MFX	1	0
5	T	4503	GOL	1	0
6	V	2101	MFX	3	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	480/481 (99%)	-0.28	4 (0%) 87 73	54, 84, 125, 146	0
1	C	480/481 (99%)	-0.31	2 (0%) 93 83	59, 82, 109, 147	0
1	R	480/481 (99%)	-0.25	13 (2%) 58 37	52, 75, 134, 176	0
1	T	480/481 (99%)	-0.27	7 (1%) 76 57	58, 87, 133, 163	0
2	B	191/193 (98%)	-0.06	8 (4%) 40 23	62, 90, 120, 129	0
2	D	190/193 (98%)	0.33	13 (6%) 20 11	80, 125, 166, 179	0
2	S	193/193 (100%)	-0.09	7 (3%) 46 28	59, 94, 125, 142	0
2	U	192/193 (99%)	0.14	12 (6%) 23 12	81, 115, 152, 163	0
3	E	20/20 (100%)	-0.38	0 100 100	63, 81, 108, 112	0
3	F	20/20 (100%)	-0.42	0 100 100	65, 86, 106, 115	0
3	V	20/20 (100%)	-0.18	1 (5%) 32 19	67, 83, 111, 123	0
3	W	20/20 (100%)	-0.37	0 100 100	63, 89, 109, 113	0
All	All	2766/2776 (99%)	-0.18	67 (2%) 62 41	52, 88, 137, 179	0

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	R	254	GLY	7.0
2	S	415	PRO	5.2
2	B	580	TYR	4.9
1	R	253	GLY	4.6
2	D	545	GLY	4.5

6.2 Non-standard residues in protein, DNA, RNA chains

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
1	PTR	T	123	16/17	0.97	0.13	-	74,78,84,87	0
1	PTR	R	123	16/17	0.97	0.13	-	73,77,90,91	0
1	PTR	C	123	16/17	0.96	0.12	-	75,79,83,86	0
1	PTR	A	123	16/17	0.95	0.14	-	80,85,91,92	0

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
5	GOL	R	4303	6/6	0.68	0.50	15.93	116,123,125,126	0
5	GOL	R	4304	6/6	0.89	0.36	9.27	97,100,100,101	0
5	GOL	A	504	6/6	0.92	0.37	5.48	85,85,86,87	0
5	GOL	C	504	6/6	0.81	0.30	2.95	97,98,99,100	0
6	MFX	F	2103	29/29	0.96	0.22	0.82	66,72,82,82	0
6	MFX	W	2103	29/29	0.94	0.23	0.57	61,71,90,91	0
6	MFX	E	2101	29/29	0.96	0.15	-0.07	61,67,75,77	0
6	MFX	V	2101	29/29	0.96	0.14	-0.43	59,68,74,75	0
6	MFX	W	2101	29/29	0.96	0.15	-0.47	76,80,91,92	0
6	MFX	F	2101	29/29	0.95	0.16	-0.51	86,90,100,101	0
4	MG	A	502	1/1	0.93	0.13	-0.83	84,84,84,84	0
4	MG	C	502	1/1	0.94	0.08	-1.46	66,66,66,66	0
4	MG	T	4502	1/1	0.94	0.08	-1.91	79,79,79,79	0
4	MG	R	4301	1/1	0.97	0.05	-2.56	73,73,73,73	0
4	MG	A	501	1/1	0.92	0.07	-3.49	76,76,76,76	0
4	MG	C	501	1/1	0.92	0.04	-3.87	94,94,94,94	0
4	MG	T	4501	1/1	0.97	0.05	-4.94	83,83,83,83	0
5	GOL	T	4503	6/6	0.80	0.32	-	92,96,98,99	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
5	GOL	C	503	6/6	0.88	0.23	-	102,103,103,103	0
4	MG	D	1001	1/1	0.98	0.10	-	66,66,66,66	0
5	GOL	A	503	6/6	0.96	0.29	-	80,80,82,83	0
4	MG	W	2104	1/1	0.76	0.49	-	75,75,75,75	0
4	MG	U	1001	1/1	0.99	0.08	-	52,52,52,52	0
4	MG	S	1001	1/1	0.99	0.12	-	52,52,52,52	0
4	MG	B	1001	1/1	0.98	0.10	-	42,42,42,42	0
4	MG	F	2102	1/1	0.97	0.05	-	69,69,69,69	0
5	GOL	R	4302	6/6	0.84	0.40	-	79,85,86,87	0
4	MG	W	2102	1/1	0.98	0.03	-	55,55,55,55	0

6.5 Other polymers (i)

There are no such residues in this entry.